

Supplementary information for example background discussion 01/07/2010

Purpose

During the December 2nd Sediment Workgroup meeting, several workgroup members said it would be useful to see an example using real data to better understand the issue of regional background. This supplementary documentation is meant to provide additional data and more in-depth discussion of the PowerPoint slide presentation that will be presented January 7th. **These examples are being presented ONLY to provide a starting point for the discussions, and do not represent an Ecology proposed approach.** There are many recognized issues regarding the use of various datasets for this purpose. We hope that despite these issues, these examples will stimulate discussion regarding the potential utility of a regional background approach, how to design studies to obtain regional data without including too much localized urban signal, alternative approaches, etc.

The “regional background” data analysis was performed on existing sediment data in Ecology’s EIM database that was collected for other purposes. Because the sampling design does not match the purpose of the data analysis, there are inherent flaws in this analysis. Decisions on what data to include were made to balance preserving enough data to do the analysis while excluding data that was most divergent from our purpose. Ideally, this analysis would be done with data from a well-designed sampling plan with sufficient recent surface sediment data, and with data-collection decisions clearly articulated. Since that is not currently available, we present this information for illustration purposes only.

Natural Background Example

In 2008, the EPA vessel *OSV Bold* provided a sampling platform for the Dredged Material Management Program (DMMP) survey of the sediments in the Puget Sound Region. The objective was to obtain sediments from around the Sound away from known sources, with the primary purpose of characterizing dioxins to assist in development of guidance for open water disposal of dredged material. Funding was obtained to actually analyze all Sediment Management Standards chemicals in addition to the planned analysis of dioxin and PCB congeners. The data from the *OSV Bold* 2008 survey was used to represent “natural background” in this example. **It should be noted that there is currently no general agreement that the “Bold” dataset represents natural background and this does not represent a decision by Ecology to use it as such.** There are additional datasets that exist that could also be used. To be consistent with existing practices, the 90th upper confidence limit (UCL) of the 90th percentile could be used as a metric. Additionally, Kaplan Meier (KM) approach for non-detects is recommended, rather than substitution of non-detects. However, since this was a preliminary example and appropriate statistical approaches for non-detects (Kaplan Meier) are time consuming, in this example non-detects were substituted with “0” and the 90th percentiles were calculated using Excel. Part of this decision was due to the fact that there were insufficient detections of Aroclors to use the KM approach, and it is not appropriate to compare summed PCB congeners from the *Bold* dataset to summed Aroclor data for other datasets.

Elliot Bay Regional Background Example

The following is provided as an example only for discussion purposes- values presented here are not part of any proposed values for the Elliot Bay region. No agreement regarding what metric to use for regional background statistical methods has been reached, and there are obvious data gaps in the existing dataset.

The most data rich embayment in the Sound, Elliot Bay, was used as a case example to determine regional background. All EIM surface data available for the bay was initially collected. The following discussion applies to Aroclor PCBs, arsenic, and mercury only.

An analysis of data trends over time (for stations that are monitored over time) indicated that older data and newer data were not statistically different. Because the age of data did not appear to skew the data (when $nd=0$), all available surface sediment data less than 20 years old were compiled, with only the most recent data being used for stations that were monitored repeatedly. Note that this analysis differed from the recent Ecology Environmental Assessment Program conclusions since no data within 250 m from shore were included in the analysis to remove any “urban signature”.

Initial attempts to determine known sources highlighted the lack of available data about the location of sources. Instead of attempting to pinpoint various sources using multiple sources of information, we decided to treat the entire shoreline as a potential source, as well as the DMMP open water disposal site and the Renton Wastewater Treatment Plant Outfall. We then filtered data to remove data within 250 m, 400 m, 800 m, and 1600 m from these sources. For the purposes of this example, it was decided to use the 400 m filter, which removed the majority of the high outliers while still leaving sufficient samples for statistical analysis.

There is no existing guidance on what metric to use for criteria based on regional background. **We have not developed specific recommendations on which statistical method or metric to use when comparing a site to “Regional Background”.** We are considering several options—a central tendency metric such as median or mean, or an upper percentile metric such as 90th percentile. There are advantages and disadvantages of each approach, which will be explored later when we review potential statistical methods. For this illustration, we have provided both a median (50th percentile) and a 90th percentile in Table 1. In the Harbor Island site example that follows, we used 50th percentile as an example for comparison of the Elliot Bay dataset to Harbor Island site data.

Summary

For the purposes of the Harbor Island site example that follows, the Elliot Bay dataset, (median value, minimum 400-meters from sources) is used to illustrate a “Regional Background” concept, and the *Bold* dataset (90th percentile, with $ND=0$) are used to illustrate a “Natural Background” concept. Total PCBs are the sum of the Aroclors. SQS and CSL numeric criteria are listed for comparison.

Table 1: Percentiles from the Elliot Bay sediment data and the *Bold* survey sediment data, as compared to Sediment Management Standards marine criteria (SQS and CSL).

Chemical	Elliot Bay 400 m from sources 50 th percentile “Regional Background” Example	Elliot Bay 400m from sources 90 th percentile “Regional Background” Example	<i>Bold</i> 90 th percentile “Natural Background” Example	SQS	CSL
Arsenic mg/kg dry weight	10.6	29.6	11.3	57	93
Mercury mg/kg dry weight	0.24	0.45	0.17	0.41	0.59
Total PCBs ug/kg dry weight	51	109	0	130	1000

Data example: Example dataset in Elliot Bay: HIRI P2 (Harbor Island).

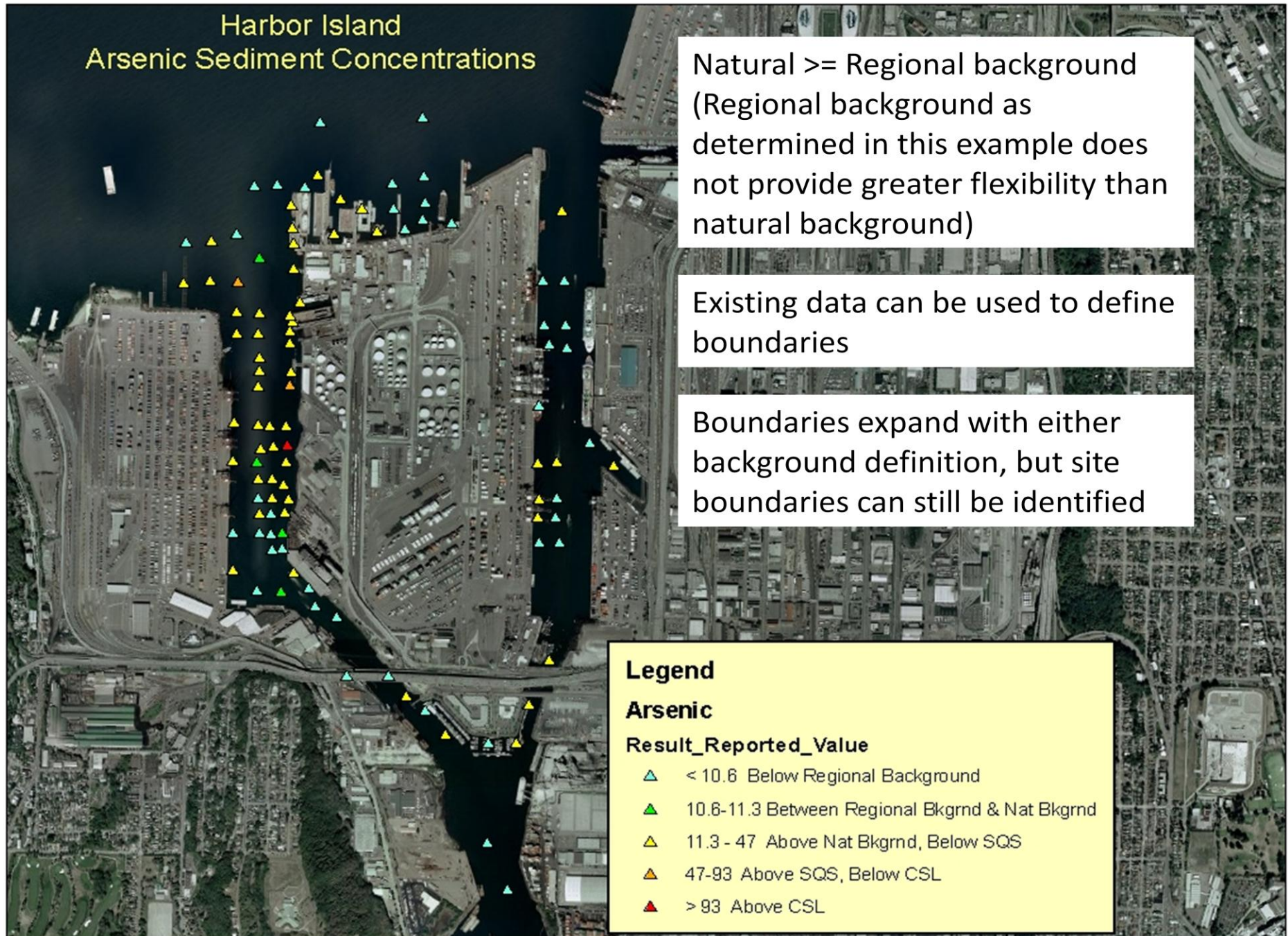
The following discussion is based on the Elliot Bay “regional” data and *Bold* dataset as “natural background” and is **for discussion purposes ONLY since there has been no agreement on use of regional background and how it might be determined.**

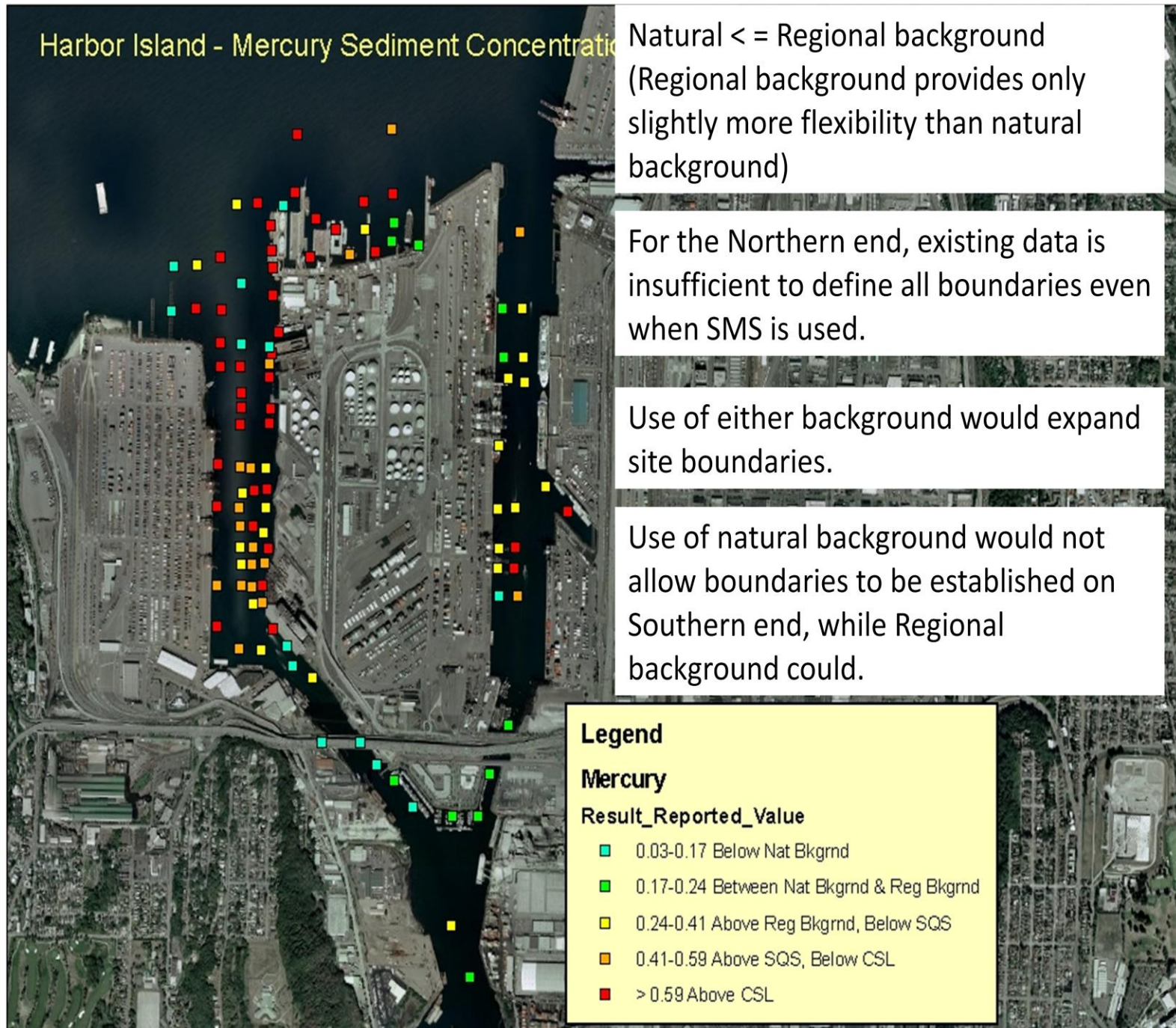
1. Data from the Harbor Island site (HIRI P2) was selected, surface samples only, replicates (field and lab) averaged. Arsenic, Mercury, and total PCBs were included in the analysis.
2. Arsenic
 - a. Only 3 of the 109 HIRIP2 samples were greater than the SQS (57 ppm), and only one of these was above the CSL (93 ppm) (all on the west/northwest side of the island).
 - b. Compared to using the SQS for identifying areas of concern, altering to the regional median (10.6 ppm) would greatly expand this site (increase to 45 of the 109 samples), and would expand to new areas of the island.
 - c. If “natural background” is defined by the *Bold* 90th percentile (11.3), impacts would not be that different than using the regional median (53 samples exceed this value).
3. Mercury
 - a. 58 of the 109 HIRI samples were greater than the SQS (0.41 ppm), and 39 of these were above the CSL (0.59 ppm).
 - b. Using the regional median (0.23) increases the number of samples exceeding “guidance” to 83 samples.
 - c. If “natural background” is defined by the *Bold* 90th percentile (0.17), areas identified as impacted would greatly expand (includes 94 of the 109 samples).
4. Total PCBs
 - a. 40 of the 109 samples exceed the SQS (130 ppb), and 5 of these exceed the CSL (1000 ppb).

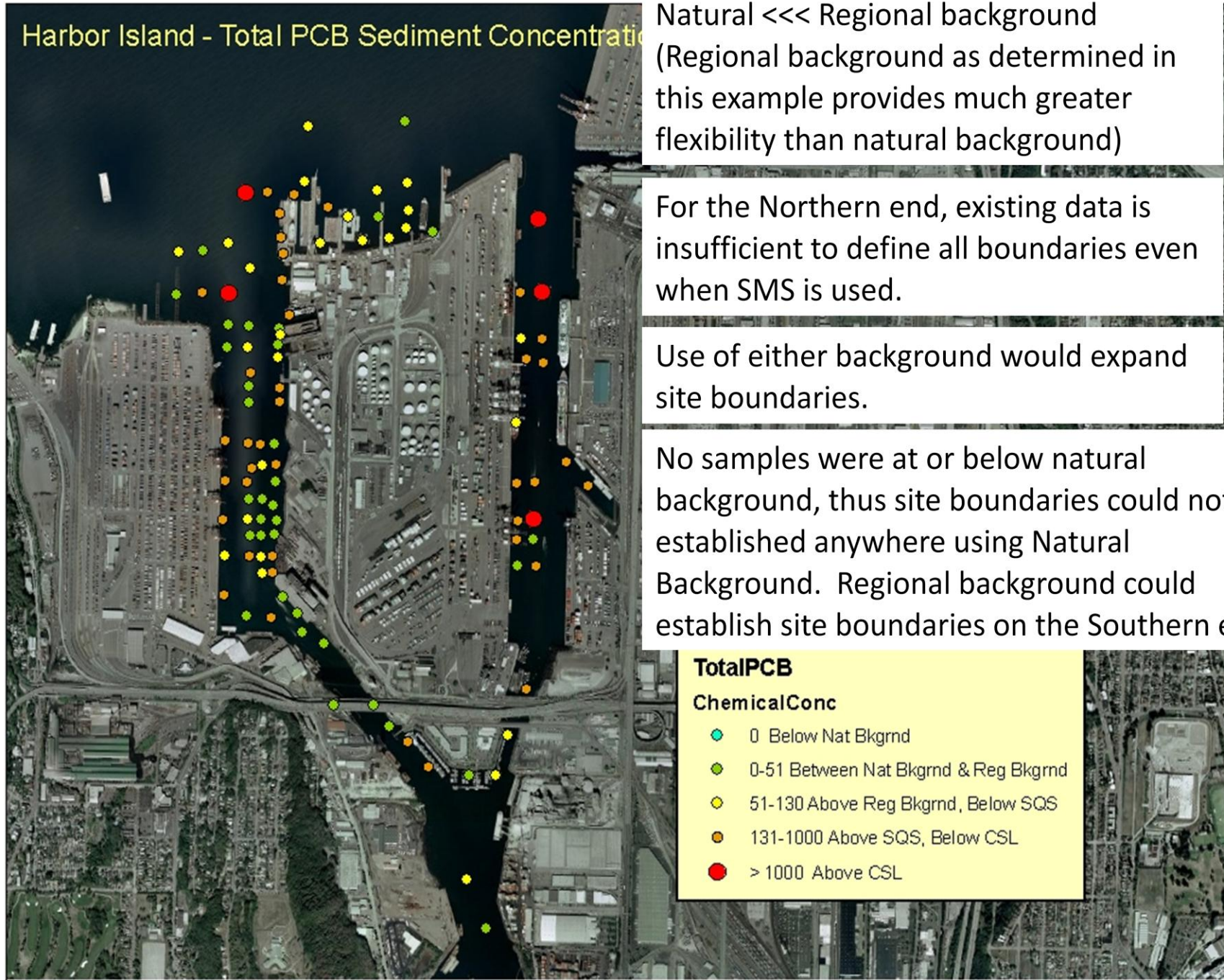
- b. If the SQS were used to define site boundaries under current guidance, then altering to the regional median (51 ppb) would greatly expand this site (67 of the 109 samples), and would expand to new areas of the island.
- c. If “natural background” is defined by the *Bold* 90th percentile, then the area identified as impacted would greatly expand, with 99 of the 109 samples exceeding the guidance.

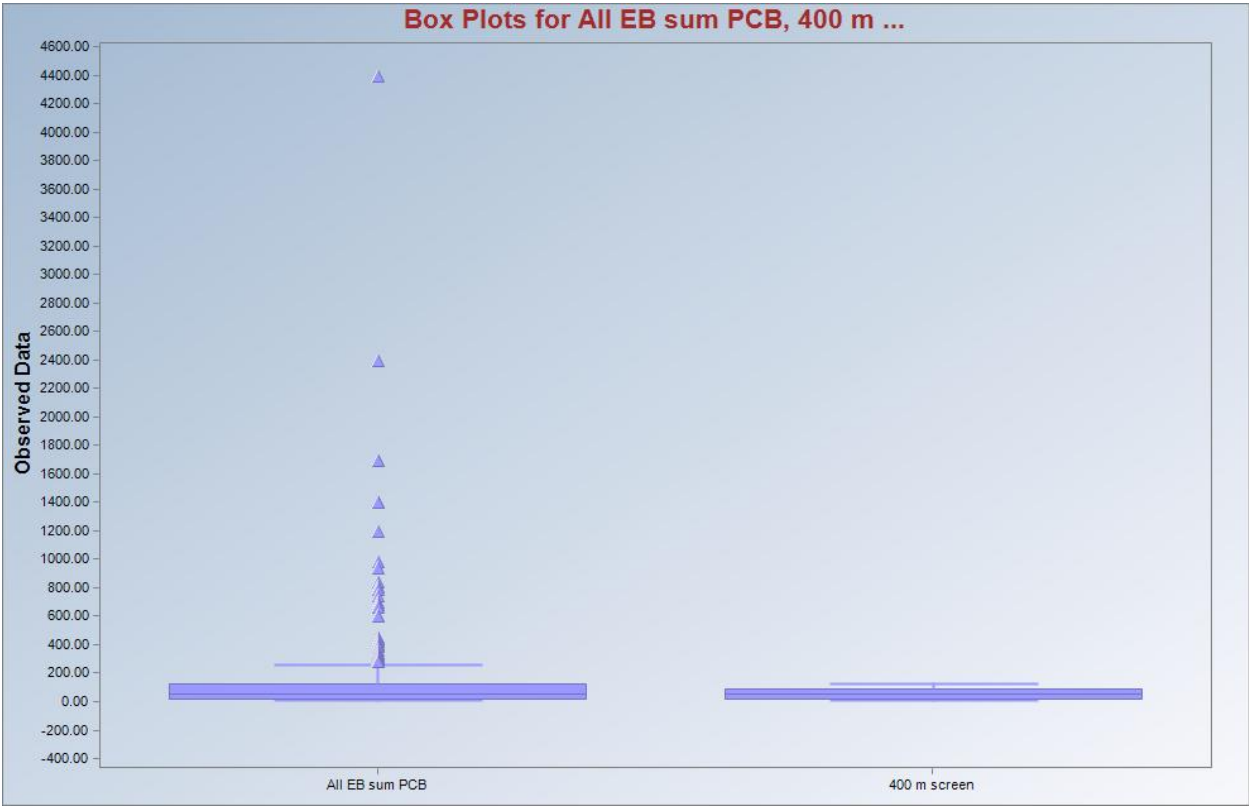
These examples cover a range of potential scenarios.

- For Arsenic, regional and natural background expands the area of concern, but the area is still defined and appears to have boundaries. Use of either definition would increase site boundaries in order to deal with bioaccumulative compounds of concern, but boundaries would be definable given the existing dataset.
- For mercury, use of regional background does not alter site boundaries based on SMS exceedances, but using natural background could result in the inability to define the site with the existing dataset, as virtually all data collected for the project exceeded natural background.
- For PCBs, even the median regional background makes it difficult to identify boundaries of the site (but this is a heavily PCB contaminated area, so no big surprise here!). Using natural background would make it nearly impossible to define the site boundaries. Without a different approach to regional background or a different approach for dealing with bioaccumulatives, it would be difficult to deal with bioaccumulative risk due to PCBs at this site.





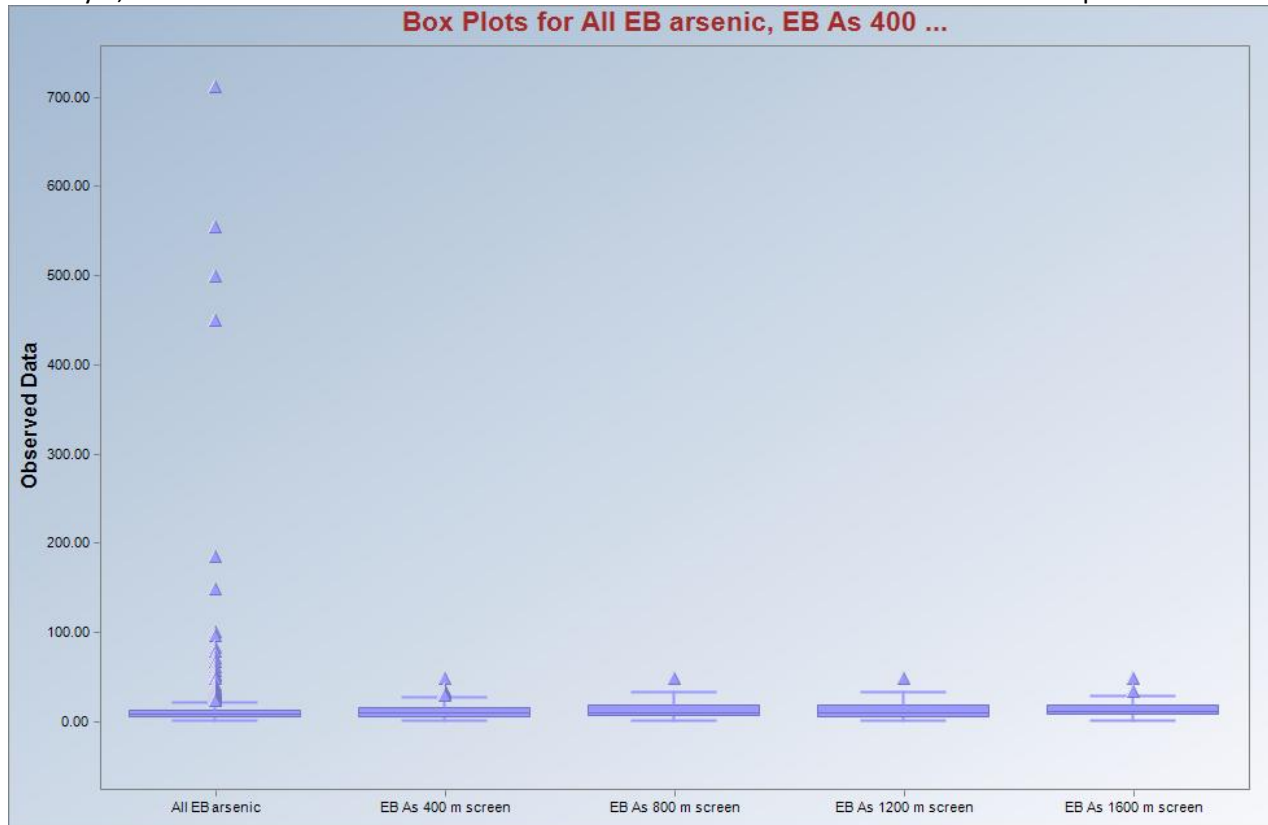




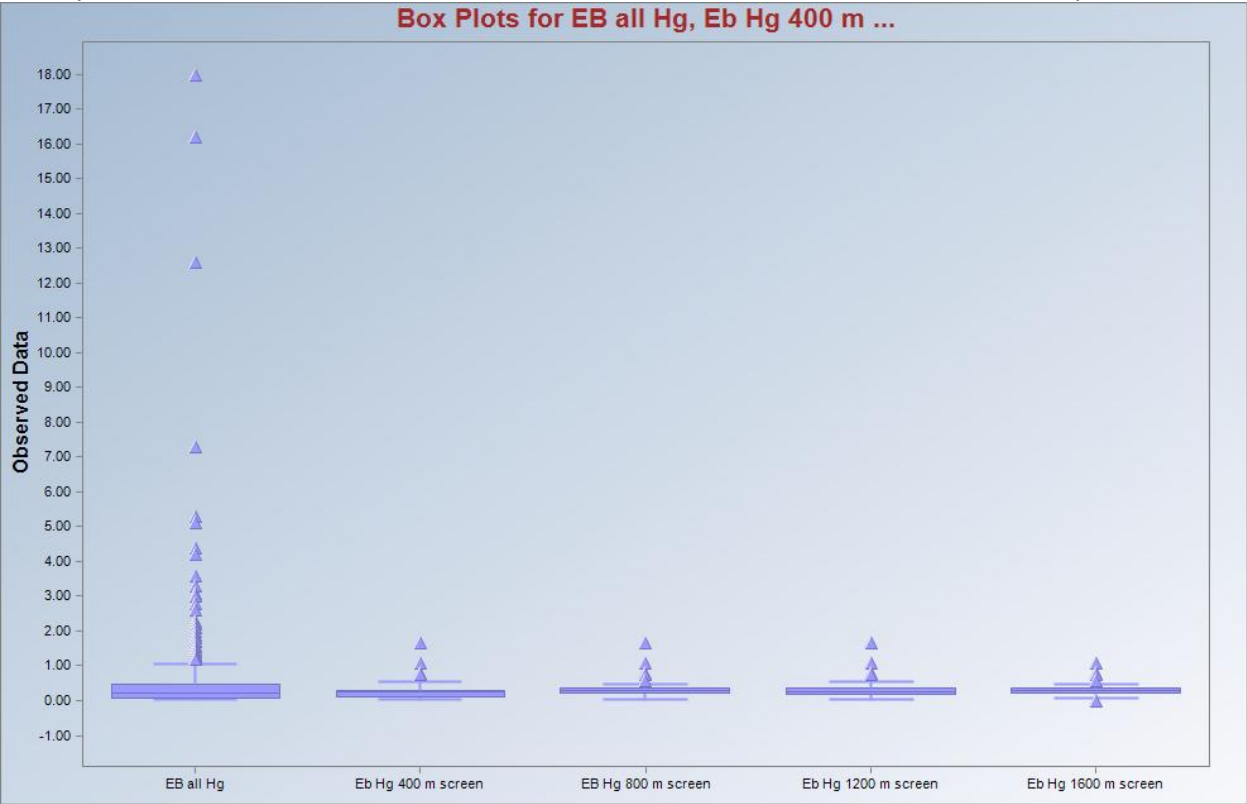
Variable		NumObs	Minimum	Maximum	Mean	Median	Variance	SD	MAD/0.675	Skewness	Kurtosis	CV
All EB sum PCB		383	1.1	4400	140.7	51	112955	336.1	48.93	7.398	75.79	2.389
400 m screen		12	0	130	56.48	51	1806	42.49	48.18	0.359	-1.102	0.752

Beyond 400 m, no samples remained.

Note that for all EB, non-detects were reported as the lowest detection limit (EIM rule).



Variable		NumObs	Minimum	Maximum	Mean	Median	Variance	SD	MAD/0.675	Skewness	Kurtosis	CV
All EB arsenic		452	0	713	17.46	8.875	2919	54.03	4.855	9.801	104.6	3.094
EB As 400 m screen		89	0	48.9	12.14	10.3	99.48	9.974	7.263	1.253	1.589	0.822
EB As 800 m screen		52	0	48.9	12.99	10.88	106	10.29	8.196	1.17	1.854	0.792
EB As 1200 m screen		52	0	48.9	12.62	10.63	111.9	10.58	8.796	1.115	1.594	0.838
EB As 1600 m screen		32	0	48.9	13.98	10.99	121.6	11.03	8.319	1.213	2.091	0.789



Variable	NumObs	Minimum	Maximum	Mean	Median	Variance	SD
EB all Hg	452	0	18	0.55	0.22	2.112	1.453
Eb Hg 400 m screen	79	0	1.69	0.259	0.239	0.0614	0.248
EB Hg 800 m screen	52	0	1.69	0.323	0.289	0.075	0.274
Eb Hg 1200 m screen	43	0	1.69	0.327	0.27	0.0901	0.3
Eb Hg 1600 m screen	28	0	1.08	0.337	0.292	0.0544	0.233